

簡易復号処理による圧縮動画データからのカット点検出

6S-9

氏原清乃 中島康之 堀裕修 加納保
KDD研究所

1 はじめに

動画の高速閲覧や効率的な検索、編集作業を可能にするための技術として動画のカット点検出があげられる。これまで、連続する動画からの検出では、DCT成分を用いてクラスタを形成し結果としてカットを検出する方法 [1]、画素変化面積の時間変化値とフィルタによる検出 [2]がある。MP EGで圧縮された動画データからの検出では後述の [3]や[4]が報告されている。前者については高精度な検出は可能なものの、各フレームのそれぞれの特徴量の抽出に時間がかかり、処理速度は再生時間の数十倍程度 [1]で、圧縮蓄積された動画からの検出にこれらを用いた場合、画像復元時間を要するため高速検出はさらに困難になる。後者については、圧縮動画データのうち離散的なフレームのデータを用いているため、高速化は実現されているものの、カット点が連続する場合や、カメラのパン、ズームといった動きの速い画像におけるカット点の未検出、過剰検出が問題となる。本稿では、[3]の方法を改良し動画圧縮データのDC成分と動き補償によりフレーム内符号化画像(I-ピクチャ)のみならずフレーム間符号化画像(P, B-ピクチャ)も簡易復号し、カット点検出のための時間的な解像度を向上させることによって、高速化を保ちつつ従来よりも高い精度で検出を行なうことが可能なことを示す。

2 従来の高速カット検出法 [3]

以下に、[3]で提案されている方式を簡単に説明する。MP EGで圧縮されたデータのうちI-ピクチャのみを用いて、フレームnのフレーム間輝度差分 D_n 、色差相関 ρ_n を求める。

ここで D_n は、

$$D_n = \sum_{i=1}^T |DY_{i,n} - DY_{i,n-1}| \quad (1)$$

で、DYは輝度のDCT-DC成分、Tはブロックの総数である。

一方 ρ_n は、

$$\rho_n = \frac{\sum_{j,k=1}^{64} (H_{n,j,k} \times H_{n-1,j,k})}{\left(\sum_{j,k=1}^{64} H_{n,j,k}^2 \times \sum_{j,k=1}^{64} H_{n-1,j,k}^2 \right)^{1/2}} \quad (2)$$

で、Hは、色差UVのDCT-DC成分を128±32の8つのカテゴリーに分け、UV平面で8×8のヒストグラム行列をとったものである。

次にカット検出の条件として以下の3つの条件を考え、いずれかに当てはまる場合カットとする。

- (1) フレーム間輝度差分の時間変化値のピーク
- (2) 色差相関の時間変化値のピーク
- (3) フレーム間輝度差分と色差相関の時間変化値のピーク

3 簡易復号化

簡易復号処理はカット点検出に必要な時空間解像度での特徴量抽出のための前処理として行う。基本的な構造としては、逆DCT処理を用いなくても可変長復号のみで抽出可能なDCT-DC成分と動き補償情報により、以下のステップでI, P, B-ピクチャを簡易復元する。図1にそのブロック図を示す。

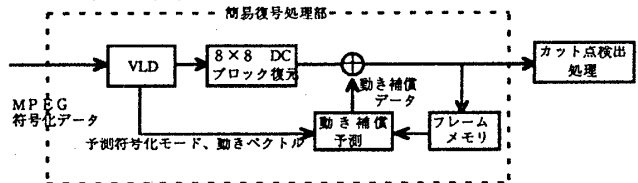


図1 簡易復号処理

(1) VLD (可変長復号部)

各ブロックデータからDCT-DC成分のみを抽出し、8×8DCブロック復元部へ入力する。さらに画面やブロック単位の子測符号化モード、動き補償モード、動きベクトルなどを復号し、これらのデータを動き補償予測部へ入力する。

(2) 8×8-DCブロック復元部

フレームを構成する各ブロックの輝度成分 $Y_{k,i,j}$ (k:輝度ブロック番号、i,j:ブロック内画素位置)、色差成分 $U_{k,i,j}$ 、 $V_{k,i,j}$ (k:色差ブロック番号、i,j:ブロック内画素位置)を次式に従って作成する。

$$Y_{k,i,j} = (DY_k) / 8 \quad i,j = 1...8 \quad (3)$$

$$U_{k,i,j} = (DU_k) / 8 \quad i,j = 1...8 \quad (4)$$

$$V_{k,i,j} = (DV_k) / 8 \quad i,j = 1...8 \quad (5)$$

A scene cut detection from compressed video using simple decoding process

Kiyono UJIHARA, Yasuyuki NAKAJIMA

Hironao HORI and Tamotsu KANO

KDD R&D Labs.

ここで、 DY_k, DU_k, DV_k は、各 8×8 ブロックのDC成分である。この処理により、膨大な処理を要する逆DCT処理を行うことなく、各フレームの簡易復号画像を得ることができる。またI-ピクチャの場合は、これが復号出力画面となる。

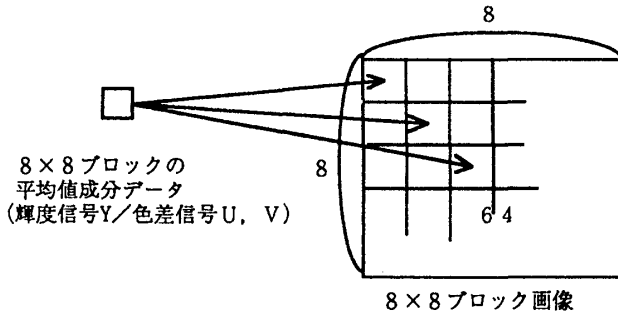


図2 8x8ブロック復元

(3) 動き補償予測部

VLDから入力された予測符号化モードが動き補償予測モードの場合は、あらかじめフレームメモリに蓄積されたI、P-ピクチャからの動き補償予測を行ない、 8×8 -DCブロック復元データを加算し最終的な復号出力画面を得る。

(4) カット点検出部

簡易復元された画像から各 8×8 ブロックの平均値を求め、前述のフレーム間輝度差分と色差相関によりカット点検出を行う。

4 実験結果

従来の検出法 [3]と簡易復号処理を用いた検出法でカット検出の比較実験を行なった。画像は、MPEG1で符号化されたニュース、料理番組、バラエティ、CMを含んだおよそ1時間のTV番組で、ワークステーション (約100MIPS)上のHDに蓄積され、そのビットストリームから検出を行なった。使用したパラメータは [3]で用いたものを採用した。

表1に検出結果を示す。ここで、方式1は従来方式でI-ピクチャのみで検出を行なったもの、方式2は全フレームを簡易復元して検出を行なったものである。

表中の%表示は、以下の計算式で求めている。

- 過(過剰検出) = (過) / カット(検出カット)
 - 未(未検出) = (未) / {(カット)-(過)+(未)}
 - 検出率 = {(カット)-(過)} / {(カット)-(過)+(未)}
- また、検出対象カットには、ワイプ、ディゾルブなどの特殊効果を使用したものも含まれている。

5 考察

表1からも分かるように、方式2ではどの画像でも検出率が向上しており、かつ方式1で目立っ

たカメラのパン、ズームといった全体的に動きの速い映像や画面内にフリップがアップになった状態で司会者がそれを差し示しているといった部分的に速い動きのある映像での過剰検出、および連続したカットでの未検出が十分低減されており、これらのカット点検出については実用レベルの高い精度を得ている。しかしNews、CMで多用されているワイプ、ディゾルブなどの特殊効果を使ったカットは、未検出となり今後の課題である。処理時間に関しては、ほぼ再生時間と同等で高速化も保たれている。

	方式1				方式2			
	カット	過-%	未-%	検出率(%)	カット	過-%	未-%	検出率(%)
News	63	8/12.7	5/8.3	91.7	57	1/1.8	4/6.7	93.3
料理	43	14/32.6	0/0	100	28	0/0	1/3.4	96.6
バラエティ	307	31/10.1	12/4.2	95.8	283	0/0	5/1.7	98.3
CM	197	16/8.1	51/22.0	78.0	212	10/4.7	30/12.9	87.1
Total	610	69/11.3	68/11.2	88.8	580	11/1.9	40/6.6	93.4

表1 検出結果

6 まとめ

本稿では、圧縮動画データからのカット検出の高精度化について述べた。動画圧縮データのDC成分と動き補償により各フレームを簡易復号し、これらのフレームを用いて検出を行なうことによって検出の時間解像度を向上させ、高速かつ高精度な検出が可能なることを確認した。日頃御指導頂くKDD研究所、浦野所長、村上次長、および古賀リーダに感謝致します。

参考文献

- [1]岩成,有木:"DCT成分を用いたシーンのクラスタリングとカット検出"信学技報,PRU93-119, 1994-01
- [2]大辻,外村:"フィルタを用いた映像カット点検出",1993信学秋季全大,D-264
- [3]中島:"フレーム間輝度差分と色差相関による圧縮動画データからのカット検出",1994信学春季全大,D-501
- [4]畑,赤間,藤木,志水:"圧縮動画データサイズを用いたカット検出法"1995信学総大,D-398